A coefficient average approximation towards Gutzwiller wavefunction formalism

Jun Liu,[1, ∗] Yongxin Yao,[1] Cai-Zhuang Wang,[1] and Kai-Ming Ho[1]

[1]*Ames Laboratory, US DOE and Department of Physics and Astronomy,*
Iowa State University, Ames, Iowa 50011, USA

(Dated: January, 26th, 2015)

Gutzwiller wavefunction is a physically well motivated trial wavefunction for describing correlated electron systems. In this work, a new approximation is introduced to facilitate evaluation of the expectation value of any operator within the Gutzwiller wavefunction formalism. The basic idea is to make use of a specially designed average over Gutzwiller wavefunction coefficients expanded in the many-body Fock space to approximate the ratio of expectation values between a Gutzwiller wavefunction and its underlying noninteracting wavefunction. To check with the standard Gutzwiller approximation (GA), we test its performance on single band systems and find quite interesting properties. On finite systems, we noticed that it gives superior performance than GA, while on infinite systems it asymptotically approaches GA. Analytic analysis together with numerical tests are provided to support this claimed asymptotical behavior. At the end, possible improvements on the approximation and its generalization towards multiband systems are illustrated and discussed.

INTRODUCTION

Correlation effects play an important role in electronic movements and physical properties of real materials[1]. Strong electron-electron interaction is believed to be key to fully understand many interesting phenomena, including high Tc superconductivity[2], heavy fermion behaviors[3], abnormal transport and optical properties[4], and more on strongly correlated materials. A versatile and convenient way to describe these systems has been posed as a big challenge to the solid state community in the past years. Part of the reasons are due to the strong Coulomb interaction preventing a controlled perturbative treatment which has been very successfully developed for weakly

---

∗Electronic address: jun.physics@gmail.com

interacting systems. The conventional mean field treatment is most versatile in gaining insights to correlated electron systems on their possible new physics, competing phases and dynamical behaviors. But its conclusions are always in question whether the ignored residual electron correlation effects are still significant enough to overshadow the presumed mean-field behaviors[5][6]. To go beyond the mean field approximation, analytic tools were proposed based on infinite summation of Feynmann diagrams of chosen types, e.g., the Random Phase Approximation (RPA)[7], the spin fluctuation theory[8], and compact diagrammatic equations like the Fluctuation Exchange approximation (FLEX)[9] or the parquet formalism[10][11]. Although they offer alternative ways for people to use, their results might be biased by starting with a subjective choice of a subgroup of diagrams which are usually numerically convenient to deal with. On the side of computational physics, a number of tools exist including Exact Diagonalization (ED)[12], Quantum Monte Carlo simulation(QMC)[13], Dynamic Mean Field Theory (DMFT) and its cluster extensions[14][15][16], and renormalization group type methods[17][18][19]. These tools are computationally demanding and can suffer from serious finite size effects or other numerical complications.

On the other hand, the variational approach[20] has always been a very important category of methods in addressing a wide range of physical problems, thanks to its capability of conveniently incorporating clear physical insights into a well-designed trial wavefunction. A typical example is the *ab initio* local density approximation (LDA) which has become an indispensable tool in modern scientific and material research[21][22]. In strongly correlated systems, the variational approach can play an important role and has been shown to be very effective and enlightening in studying various correlation effects[23][24][25]. It avoids looking for a small quantity to validate a perturbative expansion of the problem, but instead, it focuses directly on most prominent physics out of correlation effects. A simple-minded way to take care of the local onsite correlation is to form a trial wavefunction with a local projection operator to control double occupancy. This brings up the famous Gutzwiller trial wavefunction to address local correlation effects[25]. Unfortunately, a direct evaluation of operator expectation values is not practical for a many-body wavefunction. How to form a new way to efficiently implement the Gutzwiller wavefunction formalism will be the focus of the current paper.

The typical model Hamiltonian to study correlation effects is the so-called single band Hubbard model[26] on a two-dimensional square lattice, expressed as

$$\hat{H} = -t \sum_{\langle i,j \rangle, \sigma} c_{i,\sigma}^{\dagger} c_{j,\sigma} + U \sum_i n_{i,\uparrow} n_{i,\downarrow} \tag{1}$$

with i, j denoting site indices and σ spins. Two energy scales, the hopping amplitude, t, between

nearest neighboring sites and the onsite Hubbard interaction, U, exist and compete with each other. The correlation effect becomes dominant in case onsite energy U is big as compared against hopping t. The ground state wavefunction is presumably well described by the Gutzwiller trial wavefunction (GWF) defined as

$$|\Psi\rangle_G = \left(\prod_i \hat{P}_{i,G} \right) |\Psi_0\rangle \tag{2}$$

with the local Gutzwiller projector $\hat{P}_{i,G}$ defined as

$$\hat{P}_G = g^{\hat{n}_\uparrow \hat{n}_\downarrow} \tag{3}$$

Here $\hat{n}_\uparrow \hat{n}_\downarrow$ is defined as the local double occupancy operator and g is the unknown Gutzwiller variational parameter, a positive number which controls the weight of electronic configurations containing double occupancy in the noninteracting wavefunction $|\Psi_0\rangle$[27][25]. Although $\hat{P}_G$ is defined purely locally, it might still capture some nonlocal physics through the 2nd order virtual hopping process in the strong correlation limit[28][29]. Thus Eq. 2 is a quite reasonable trial wavefunction to capture the essential physics of a correlated electron system.

However, introducing a physically sensible trial wavefunction is only the first step. Brute force evaluation of physical observables in a many-body system can be a real pain due to numerical difficulties and computer hardware capacity. Thus, the so-called Gutzwiller approximation (GA) was introduced to facilitate the evaluation[25][30][31][32]. There, the expectation value for an interacting system is assumed to be proportional to that of the noninteracting wavefunction through a site-decomposeable renormalization factor. Taking the hopping operator as an example, GA gives

$$\left\langle c^\dagger_{i,\alpha,\sigma} c_{j,\beta,\sigma} \right\rangle = z_{i,\alpha} z_{j,\beta} \left\langle c^\dagger_{i,\alpha,\sigma} c_{j,\beta,\sigma} \right\rangle_0 \tag{4}$$

for two distinct sites. Here, α, β denote orbital indices and σ denotes spin. A succinct analytic expression for $z_{i,\alpha}$ is available for the operator $c^\dagger_{i,\alpha,\sigma}$ (or $c_{i,\alpha,\sigma}$)[32]. Actually, we might come up with an educated guess that, given any local operator $\hat{o}_I$ with I the set of local spin-orbital states involved in $\hat{o}$, the corresponding $z_{\hat{o}}$ is given as

$$z_{\tilde{o}} = \frac{1}{\sqrt{\prod_{(\alpha,\sigma) \in I} n^0_{(\alpha,\sigma)} \left(1 - n^0_{(\alpha,\sigma)}\right) A_{(\alpha,\sigma)}}} \left(\sum_{\Gamma,\Gamma'} \sqrt{p_\Gamma p_{\Gamma'}} \left| \langle \Gamma' | \hat{o}_I | \Gamma \rangle \right|^2 \right) \tag{5}$$

with

$$A_{(\alpha,\sigma)} = \begin{cases} \dfrac{n^0_{(\alpha,\sigma)}}{1 - n^0_{(\alpha,\sigma)}} & \text{if } \hat{n}_{(\alpha,\sigma)} \text{ is part of } \tilde{o}_I \\ 1 & \text{o.w.} \end{cases} \tag{6}$$

Here $n^0_{(\alpha,\sigma)} = \langle\Psi_0|\,\hat{n}_{(\alpha,\sigma)}\,|\Psi_0\rangle$ and p_Γ denotes the occupational probability at a local Fock state Γ. The summation in Eq. 5 traverses all local Fock states mutually related by the local operator $\hat{o}_I$. One can go through the tedious yet well-established procedure to verify this relations for any operator[33]. The introduction of $z_{i,\alpha}$ as a renormalization factor is physically very sensible. Strong correlation effects among electrons necessarily leave an impact on their motions, thus modifying their dynamical and transport behaviors as compared against a Fermi liquid system described by $|\Psi_0\rangle$.

There have been roughly three different ways to reach GA. Metzner et al approached GA with mathematical rigor by applying Feymann diagrammatic expansion techinque to the Gutzwiller wavefunction[34][35]. The approach effectively cuts intersite communications in the infinite spatial dimension limit and leaves only intrasite correlations among local orbits. This brings up the main site-wise decomposibility feature of GA readily seen from the site-dependent renormalization factors illustrated in Eq. 4. This approach has been further carried out by Bunemann et. al. and Fabrizio et. al. to multiband systems and with a more generalized Gutzwiller projector[32][36] [37][38]. The second way of writing down GA is through physical intuition with a hand-waving argument. Gutzwiller originally formulated GA based on some assumptions[25] whose physics were rather obscure thus preventing it from being generalized to multiband systems. Ogawa's counting argument formally pointed out that the physics underlying Gutzwiller's GA is to assume that the expectation value of a product of number operators equals the product of expectation value of each number operator in the series[31]. The physics of GA is most transparent in Bunemann's version of counting argument[32]. We regard it as a third approach towards GA due to its balance between clarity in physics and completeness in formulation. Bunemann made it clear that the projection of a noninteracting wavefunction onto a specific electronic configuration depends only on the number of electrons on each local orbital on the lattice, but not on how these electrons are distributed on the lattice. This implies two things. First, the combinatoric trick accompanying the assumption can be used to evaluate any inner product, which greatly simplifies the whole formalism and introduces an additional convenience of taking the thermodynamic limit to further simplify the expressions. Second, he pointed out that GA actually implies more than just one relation, but instead, infinitely many for different operators whose expectation values are to be evaluated. This leaves us a taste on how crude GA might work in reality.

Despite many achievements made by GA on qualitatively addressing the Mott physics in correlated electron systems, quantitatively however, it might introduce artifacts. For instance, GA predicts existence of a Brinkman-Rice metal-insulator transition at a critical onsite interaction

comparable to the bandwidth in a single band system[39], while GWF formalism should always give metallic behavior unless $U \to \infty$[34]. These artifacts are of course closely related with the rough assumptions made in GA. The artificial Brinkman-Rice transition seems to imply excessive local correlations were introduced by GA. In addition, besides the unphysical infinite spatial dimension limit, the most questionable assumption in GA is that it calls for a presumably known connection between the Gutzwiller local orbital occupations and those of the noninteracting wavefunction, a requisite to facilitate the Feynman diagrammatic expansion to derive GA[32][33]. This might cause trouble in multiband systems where charge flow among orbitals might necessarily play a role due to correlation effects[40].

It would be useful if some of the concerns mentioned above could be addressed. However, within the existing GA formulation, this is not an easy task mainly because the formulation is mathematically prohibitively complicated to be further improved. Thus we need a new way to look into the renormalization factors of operators, central quantity of GA to bridge between the Gutzwiller and the noninteracting wavefunction. We come up with a very simple way to design the renormalization factors which gives the standard GA in the thermodynamic limit, and which also has the potential to be improved systematically. We might thus call it a fourth way to reach GA.

METHOD

Let's introduce some notations first to facilitate the up-coming discussion. Given a number of local spin-orbital states, the local Fock space is set up and denoted as Γ. Specifically, in the single band case, we denote the spin-orbital states as $1 = (\alpha, \uparrow)$, $2 = (\alpha, \downarrow)$ with α the orbital index, and note $\Gamma \in \{\emptyset, (1), (2), (1, 2)\}$. Here $\emptyset$ denotes an empty state and (1) denotes a singly occupied state with spin-orbital state 1 taken. Similar interpretation applies to (2) and $(1, 2)$. Let's denote $|\Gamma|$ to represent the number of electrons in a Fock state Γ, and $s_z(\Gamma)$ the total spin z component of Γ. Specifically, we have the following

$$|\emptyset| = 0, s_z(\emptyset) = 0;$$

$$|(1)| = 1, s_z(1) = \frac{1}{2};$$

$$|(2)| = 1, s_z(2) = -\frac{1}{2};$$

$$|(1, 2)| = 2, s_z(1, 2) = 0$$

Suppose the system has a total of N sites, N_e electrons and S_z net spins. Each electronic configuration is described in an occupation representation as $\{\boldsymbol{\Gamma}\} = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_N\}$ with subscripts denoting site indices. These occupation representations form a complete Fock space on the lattice, satisfying

$$\sum_{\{\boldsymbol{\Gamma}\}} |\boldsymbol{\Gamma}\rangle \langle \boldsymbol{\Gamma}| = I \tag{7}$$

The count of each local Fock state in $\boldsymbol{\Gamma}$, $n_\Gamma(\boldsymbol{\Gamma})$, is defined to be

$$n_\Gamma(\boldsymbol{\Gamma}) = \sum_i n_{i,\Gamma}(\Gamma) \text{ with } n_{i,\Gamma}(\Gamma) = \delta_{\Gamma,\Gamma_i} \tag{8}$$

or simply n_Γ if no confusion arouses in its interpretation. All possible values for n_Γ form a set denoted as $\{n_\Gamma\}$.

Here is something quite general and useful for us to know. Consider a generic trial wavefunction $|\Psi\rangle$ and note the closure relation of Eq. 7, the inner product involving any operator $\hat{O}$ is written as

$$\langle \Psi | \hat{O} | \Psi \rangle = \sum_{\{\boldsymbol{\Gamma}\},\{\boldsymbol{\Gamma}'\}} \langle \Psi | \boldsymbol{\Gamma} \rangle \langle \boldsymbol{\Gamma} | \hat{O} | \boldsymbol{\Gamma}' \rangle \langle \boldsymbol{\Gamma}' | \Psi \rangle \tag{9}$$

This can be reexpressed in terms of a noninteracting wavefunction $|\Psi_0\rangle$ on which $|\Psi\rangle$ is set up as

$$\langle \Psi | \hat{O} | \Psi \rangle = \sum_{\{\boldsymbol{\Gamma}\},\{\boldsymbol{\Gamma}'\}} C(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}') \langle \Psi_0 | \boldsymbol{\Gamma} \rangle \langle \boldsymbol{\Gamma} | \hat{O} | \boldsymbol{\Gamma}' \rangle \langle \boldsymbol{\Gamma}' | \Psi_0 \rangle \tag{10}$$

with

$$C(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}') = \frac{\langle \Psi | \boldsymbol{\Gamma} \rangle}{\langle \Psi_0 | \boldsymbol{\Gamma} \rangle} \frac{\langle \boldsymbol{\Gamma}' | \Psi \rangle}{\langle \boldsymbol{\Gamma}' | \Psi_0 \rangle} \tag{11}$$

which holds as long as the denominators do not vanish, a fairly mild constraint to be met in most cases. If we manage to replace $C(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}')$ with some constant $\bar{C}$, an idea borrowed from the first mean value theorem of integrals, Eq. 10 is now

$$\langle \Psi | \hat{O} | \Psi \rangle \simeq \bar{C} \sum_{\{\boldsymbol{\Gamma}\},\{\boldsymbol{\Gamma}'\}} \langle \Psi_0 | \boldsymbol{\Gamma} \rangle \langle \boldsymbol{\Gamma} | \hat{O} | \boldsymbol{\Gamma}' \rangle \langle \boldsymbol{\Gamma}' | \Psi_0 \rangle$$
$$= \bar{C} \langle \Psi_0 | \hat{O} | \Psi_0 \rangle \tag{12}$$

In the like fashion, the expectation value of $\hat{O}$ can be related to that of the noninteracting system through

$$\left\langle \hat{O} \right\rangle = \frac{\langle \Psi | \hat{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \simeq \frac{\mathcal{F}}{\mathcal{B}} \frac{\langle \Psi_0 | \hat{O} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} = Z_{\hat{O}} \left\langle \hat{O} \right\rangle_0 \tag{13}$$

with

$$Z_{\hat{O}} = \frac{\mathcal{F}}{\mathcal{B}} \qquad (14)$$

Here $\mathcal{F}$ and $\mathcal{B}$ are specifically chosen symbols to denote the constant prefactors in the numerator and the denominator respectively. Just like GA[32], we call $Z_{\hat{O}}$ the renormalization factor for operator $\hat{O}$. Thus, through a renormalization factor, the expectation value of an operator evaluated with a correlated Gutzwiller wavefunction is directly related to that of a noninteracting wavefunction. In principle, Eq. 13 is able to rigorously hold if $\mathcal{F}$ and $\mathcal{B}$ are chosen correctly. However, this is most likely not the case in practice. How well Eq. 13 holds depends on the specific method used to set up $Z_{\hat{O}}$.

We believe the implication of the form of a renormalization factor in Eq. 13 reaches far beyond GA in that it gives us a clear route to determine $Z_{\hat{O}}$ without resorting to the very involved algebras from Feynmann diagrams and the unphysical assumptions used in GA. One way to determine $Z_{\hat{O}}$, or equivalently $\mathcal{F}$ and $\mathcal{B}$, is to go through $C\left(\mathbf{\Gamma}, \mathbf{\Gamma}'\right)$ directly by averaging these wavefunction coefficients in the many-body Fock space. You will see that, in the single band case, even a simple definition towards $\mathcal{F}$ and $\mathcal{B}$ through a wavefunction coefficient average readily gives GA as its limiting behavior. Here are the details for defining $\mathcal{F}$ and $\mathcal{B}$.

For the Gutzwiller trial wavefunction defined in Eq. 2 and Eq. 3 with only density operators entering the Gutzwiller projector, one possible way to define $\bar{C}$ for any operator $\hat{O}$ is

$$\bar{C}_{\hat{O}}\left(g\right) = \frac{\sum_{\{\mathbf{\Gamma}\},\{\mathbf{\Gamma}'\}} \left[\prod_i g^{n_{i,\Gamma_D}(\mathbf{\Gamma})+n_{i,\Gamma_D}(\mathbf{\Gamma}')}\right] \mathcal{S}\left(\hat{O};\mathbf{\Gamma},\mathbf{\Gamma}'\right) \delta_{\mathbf{\Gamma}}^{(3)}\delta_{\mathbf{\Gamma}'}^{(3)}}{\sum_{\{\mathbf{\Gamma}\},\{\mathbf{\Gamma}'\}} \mathcal{S}\left(\hat{O};\mathbf{\Gamma},\mathbf{\Gamma}'\right) \delta_{\mathbf{\Gamma}}^{(3)}\delta_{\mathbf{\Gamma}'}^{(3)}} \qquad (15)$$

where $n_{i,\Gamma_D}\left(\mathbf{\Gamma}\right) = \delta_{\Gamma_D,\Gamma_i}$ with $\Gamma_D = c_\uparrow^\dagger c_\downarrow^\dagger |0\rangle$ denoting the doubly occupied Fock state. The delta functions,

$$\delta_{\mathbf{\Gamma}}^{(3)} = \delta\left(\sum_\Gamma n_\Gamma - N\right) \delta\left(\sum_\Gamma n_\Gamma |\Gamma| - N_e\right) \delta\left(\sum_\Gamma n_\Gamma s_z\left(\Gamma\right) - S_Z\right) \qquad (16)$$

ensure conservation of total charge and spin, and expansion on all sites. The indicator function $\mathcal{S}\left(\hat{O};\mathbf{\Gamma},\mathbf{\Gamma}'\right)$ is defined as

$$\mathcal{S}\left(\hat{O};\mathbf{\Gamma},\mathbf{\Gamma}'\right) = \begin{cases} 1 & \text{if } \langle\mathbf{\Gamma}|\,\hat{O}\,|\mathbf{\Gamma}'\rangle \neq 0 \\ 0 & \text{o.w.} \end{cases} \qquad (17)$$

Thus the operator $\hat{O}$ has its effect included through $\mathcal{S}\left(\hat{O};\mathbf{\Gamma},\mathbf{\Gamma}'\right)$. Although looked complicated, Eq. 15 is just a carefully designed average over nonvanishing elements of the inner product $_G\langle\Psi|\,O\,|\Psi\rangle_G$

expanded in the Fock space set up on the whole lattice. Thus it readily guarantees the correct limiting behavior as g approaches unity. Why $\bar{C}_{\hat{O}}$ can be conveniently written out in Eq. 15 is closely related to the current choice of the Gutzwiller projector defined on number operators only. An additional advantage with such a Gutzwiller projector is, this enables us to use combinatorics to simplify the whole expression.

Now we are ready to define $\left\langle \hat{O} \right\rangle$ and its renormalization factor $Z_{\hat{O}}$. Without loss of generality, let's assume $\hat{O} = \hat{O}_{ijkl}$ acts on a sublattice $\mathcal{R} = \{i, j, k, l\}$ defined on the affected sites, and denote the number of distinct sites in $\mathcal{R}$ to be $N_{\mathcal{R}}$. We define $\tilde{\mathcal{R}}$ as the complementary lattice to $\mathcal{R}$ and denote its number of sites to be $N_{\tilde{\mathcal{R}}}$. A Fock space on the whole lattice can be decomposed into Fock spaces on the sublattices $\mathcal{R}$ and $\tilde{\mathcal{R}}$. The counts on local Fock states in both sublattices are denoted as n_{Γ} and $\tilde{n}_{\Gamma}$ respectively. Then, apply Eq. 15 to both numerator and denominator and we have

$$\left\langle \hat{O}_{ijkl} \right\rangle \simeq \frac{\mathcal{F}}{\mathcal{B}} \left\langle \hat{O}_{ijkl} \right\rangle_0 \tag{18}$$

with

$$\mathcal{F} = \bar{C}_{\hat{O}}(g) = \frac{\sum'_{\{\Gamma\},\{\Gamma'\}} \left\{ \left(\prod_{i \in \mathcal{R}} g^{n_{i,\Gamma_D}(\Gamma) + n_{i,\Gamma_D}(\Gamma')} \right) \mathcal{S}\left(\hat{O}; \Gamma, \Gamma'\right) \left[\sum''_{\{\tilde{n}_{\Gamma}\}} \mathcal{C}_{N_{\tilde{\mathcal{R}}}}^{\{\tilde{n}_{\Gamma}\}} g^{2\tilde{n}_{\Gamma_D}} \tilde{\delta}^{(3)} \right] \right\}}{\sum'_{\{\Gamma\},\{\Gamma'\}} \left\{ \mathcal{S}\left(\hat{O}; \Gamma, \Gamma'\right) \left[\sum''_{\{\tilde{n}_{\Gamma}\}} \mathcal{C}_{N_{\tilde{\mathcal{R}}}}^{\{\tilde{n}_{\Gamma}\}} \tilde{\delta}^{(3)} \right] \right\}} \tag{19}$$

$$\mathcal{B} = \bar{C}_{\hat{I}}(g) = \frac{\sum_{\{n_{\Gamma}\}} \left\{ \mathcal{C}_{N}^{\{n_{\Gamma}\}} g^{2n_{\Gamma_D}} \delta_{\Gamma}^{(3)} \right\}}{\sum_{\{n_{\Gamma}\}} \left\{ \mathcal{C}_{N}^{\{n_{\Gamma}\}} \delta_{\Gamma}^{(3)} \right\}} \tag{20}$$

Here $\hat{I}$ denotes the identity operator. The primed sum in Eq. 19 denotes that $\{\Gamma\}$ and $\{\Gamma'\}$ are confined within sublattice $\mathcal{R}$ only. The double primed sum assumes $\tilde{n}_{\Gamma}$ and $\tilde{n}_{\Gamma_D}$ are counted within sublattice $\tilde{R}$ for fixed Γ. The delta functions $\tilde{\delta}^{(3)}$ states the conservation constraints reenforced on $\tilde{\mathcal{R}}$ are

$$\tilde{\delta}^{(3)} = \delta\left(\sum_{\Gamma}'' \tilde{n}_{\Gamma} - N_{\tilde{N}} \right) \delta\left(\sum_{\Gamma}'' \tilde{n}_{\Gamma} |\Gamma| - \tilde{N}_e \right) \delta\left(\sum_{\Gamma}'' \tilde{n}_{\Gamma} s_z(\Gamma) - \tilde{S}_Z \right) \tag{21}$$

with

$$\tilde{N}_e = N_e - N_e^{\mathcal{R}}(\Gamma) \tag{22}$$

$$\tilde{S}_Z = S_Z - S_Z^{\mathcal{R}}(\Gamma) \tag{23}$$

where $N_e^{\mathcal{R}}(\Gamma)$ and $S_Z^{\mathcal{R}}(\Gamma)$ denote total number of electrons and net spins on sublattice $\mathcal{R}$ occupied by a specific electronic configuration Γ. Note, a physical operator would not change electron

occupation and total spin z component, which implies $N_e^{\mathcal{R}}(\boldsymbol{\Gamma}) = N_e^{\mathcal{R}}(\boldsymbol{\Gamma}')$ and $S_Z^{\mathcal{R}}(\boldsymbol{\Gamma}) = S_Z^{\mathcal{R}}(\boldsymbol{\Gamma}')$. $\mathcal{C}_N^{\{n_\Gamma\}}$ takes the standard definition of the multinomial coefficient,

$$\mathcal{C}_N^{\{n_\Gamma\}} = \frac{N!}{\prod_\Gamma (n_\Gamma!)} \tag{24}$$

Eq. 18 to Eq. 23 completes the expressions for approximating the renormalization factors through wavefunction coefficient averaging. These expressions can be directly used to approximate the energy expectation value of a Hamiltonian of a small system.

For big systems, one needs to take the thermodynamic limit to simplify the coefficients in Eq. 18. A typical sum involved in Eq. 19 and Eq. 20 is

$$A(\{\tilde{n}_\Gamma\}) = \sum_{\{\tilde{n}_\Gamma\}}^{''} \mathcal{C}_{N_{\tilde{\mathcal{R}}}}^{\{\tilde{n}_\Gamma\}} g^{2\tilde{n}_{\Gamma_D}} \delta^{(3)} \tag{25}$$

In the thermodynamic limit, the sum can be replaced by a single term located at some unknown $\{\tilde{n}_\Gamma^*\}$ where the terms are peaked. To solve for $\{\tilde{n}_\Gamma^*\}$ we introduce Lagrange multipliers to relax the constraints entering Eq. 25 and replace the factorials with their asymptotic expressions using the Sterling's formula. We finally come up with the following functional

$$h(\{\tilde{n}_\Gamma\}, \alpha, \beta, \gamma) = \tilde{N} \ln \tilde{N} - \sum_\Gamma \tilde{n}_\Gamma \ln \tilde{n}_\Gamma + \sum_\Gamma 2\tilde{n}_\Gamma \ln g_\Gamma$$
$$+ \alpha \left(\sum_\Gamma{}' \tilde{n}_\Gamma - \tilde{N} \right) + \beta \left(\sum_\Gamma{}' \tilde{n}_\Gamma |\Gamma| - \tilde{N}_e \right) + \gamma \left(\sum_\Gamma{}' \tilde{n}_\Gamma s_z(\Gamma) - \tilde{S}_z \right) \tag{26}$$

whose minimization gives $\{\tilde{n}_\Gamma^*\}$. The Gutzwiller variational parameter, g, now carries a local Fock state index for the purpose of formalism consistency. Following the standard procedure to extremize $h(\{\tilde{n}_\Gamma\}, \alpha, \beta, \gamma)$ by taking partial derivatives with respect to $\{\tilde{n}_\Gamma\}$, we get explicit expressions for $\tilde{n}_\Gamma^*$ as

$$\frac{\tilde{n}_\Gamma^*}{\tilde{N}} = \frac{g_\Gamma^2 \exp(\beta |\Gamma| + \gamma s_z(\Gamma))}{\sum_\Gamma g_\Gamma^2 \exp(\beta |\Gamma| + \gamma s_z(\Gamma))} \tag{27}$$

The remaining unknown, β and γ, satisfy the following two nonlinear equations,

$$\sum_\Gamma \left(|\Gamma| - \tilde{N}_e/\tilde{N} \right) g_\Gamma^2 \exp(\beta |\Gamma| + \gamma s_z(\Gamma)) = 0 \tag{28}$$

$$\sum_\Gamma \left(s_z(\Gamma) - \tilde{S}_z/\tilde{N} \right) g_\Gamma^2 \exp(\beta |\Gamma| + \gamma s_z(\Gamma)) = 0 \tag{29}$$

Note $\tilde{N}_e$ and $\tilde{S}_z$ defined in Eq. 22 and Eq. 23 explicitly enter the above two equations. This implies that each term of the primed sums in $\mathcal{F}$ and $\mathcal{B}$ will have its own optimal $\{\tilde{n}_\Gamma^*\}$. This complicates

the whole computation, but is expected to have a mild impact on its performance as Eq. 28 and Eq. 29 are usually well-behaved.

Eq. 18 and its accompanying definition of Eq. 19 and Eq. 20 looks fundamentally different from GA in both the underlying assumptions and their mathematical formulations. Surprisingly, however, the current scheme can be shown to be well related to GA in the single band system. A brief proof is provided in Appendix A and a numerical study is given in Fig. 4 in the main text. It clearly reveals that GA can be regarded as being an overly simplified approximation of the current scheme, but with vanishing difference at infinite lattice size. Thus, we have provided here a fourth perspective to look into GA and its local nature accompanying the infinite spatial dimension limit towards the GWF formalism. Furthermore, the proof readily suggests that such a simple scheme has superior performance than GA in finite systems. This fact is case studied here on small Hydrogen clusters. To distinguish it from GA, we call it AA in the coming discussions.

COMPARISON BETWEEN THE NEW APPROXIMATION AND GA

Let's first look at the performance of AA and GA on predicting the ground state energy of a finite system constructed by Hydrogen atoms. We choose the one-dimensional minimum basis Hydrogen chains with periodic boundary conditions to carry out the calculation. The *ab initio* Hamiltonian is

$$H = \sum_{(i,j),\sigma} t_{ij} c_{i,\sigma}^\dagger c_{j,\sigma} + \sum_{(i,j,k,l);\sigma,\sigma'} U_{i,j,k,l} c_{i,\sigma}^\dagger c_{j,\sigma'}^\dagger c_{k,\sigma'} c_{l,\sigma} \tag{30}$$

where (i,j) runs through all possible site pairs and (i,j,k,l) describes all possible 2-body interactions on the chain. The bare energy parameters t and U are evaluated with GAMESS (US), a Quantum Chemistry package widely used for molecular calculations. Among all the interactions, the density-density interactions are the most dominant. Its ratio against nearest neighbor hopping is a quantitative measure of the strength of correlation, which increases as inter-atomic distance increases. The Gutzwiller wavefunction is constructed by applying the Gutzwiller projector defined in Eq. 3 onto the noninteracting Hartree-Fock wavefunction formed by the occupied molecular orbitals. We check the performance of both approximations on two systems formed by 8 and 16 hydrogen atoms respectively. The brute force evaluation with the Gutzwiller trial wavefunction (GWF) and the Configuration Interaction (CI) calculation give benchmark results for AA and GA to compare. As shown in Fig. 1, GWF gives quite close energy to CI and is able to reach the correct atomic energy given by CI. One might note that CI gives an atomic Hydrogen ground state
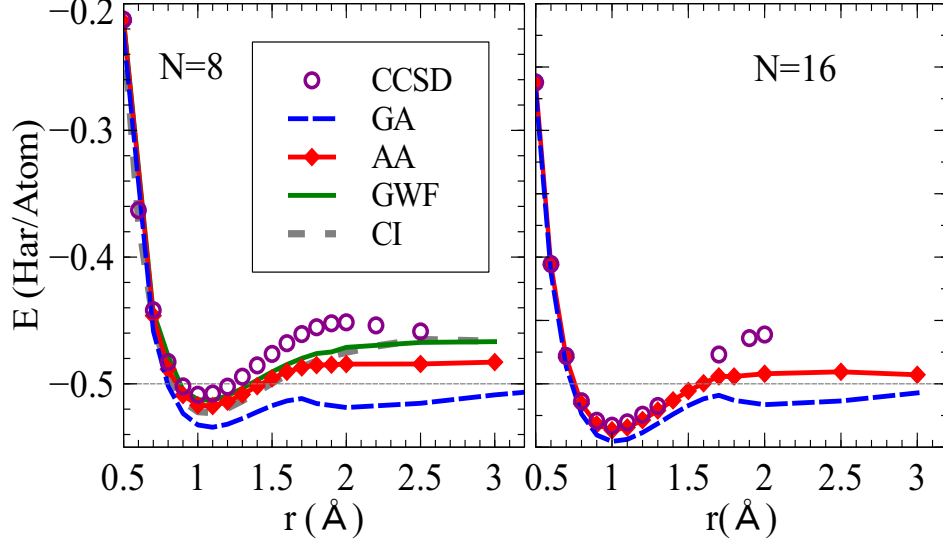
FIG. 1: Total energy evaluated using the current method (denoted as AA), GA, Coupled Cluster (CCSD), Configuration Interaction (CI), and GWF if possible, on circular Hydrogen (H) chains of N=8 and N=16 atoms. The *ab initio* Hamiltonians are returned by GAMESS(US) using a STO-3G basis set description of H. For N=8, GWF and CI results are both available to benchmark AA and GA. For N=16, only CCSD is evaluated to check both approximations. Note CCSD does not converge well beyond the bond breaking region. GWF and CI do not reach the known Hydrogen atomic energy of -0.5Hartree due to the use of minimum basis set. The noninteracting wavefunction underlying GWF (thus GA and AA as well) are fixed at the Hartree-Fock solution.

energy differing from the well-known value of $-13.6eV$. This is due to the STO-3G minimum basis set chosen to set up the *ab initio* Hamiltonian in Eq. 30. In the 16 Hydrogen atom system, GWF and CI are not convenient to evaluate. Thus, we use the Coupled Cluster method(CCSD) instead to give an estimate on the ground state energy. CCSD is a very popular method in Quantum Chemistry[41] with good balance between speed and accuracy. It has some known issues but is good enough for our current purpose. The comparison between AA and GA, together with other methods, are presented in Fig. 1, where the ground state energies are plotted against different inter-atomic separations. Several things are quite interesting to note. GA is in general not as good as AA, but its performance seems to be improved as system size increases. This gives a strong support on our proof of GA being a grand canonical ensemble description of the current coefficient averaging approximation, while AA is its canonical ensemble description. Another thing to notice is that, as compared with GWF, AA is slightly higher in energy around the bonding region and does not reach the atomic limit energy returned by GWF. This is mainly due to biased
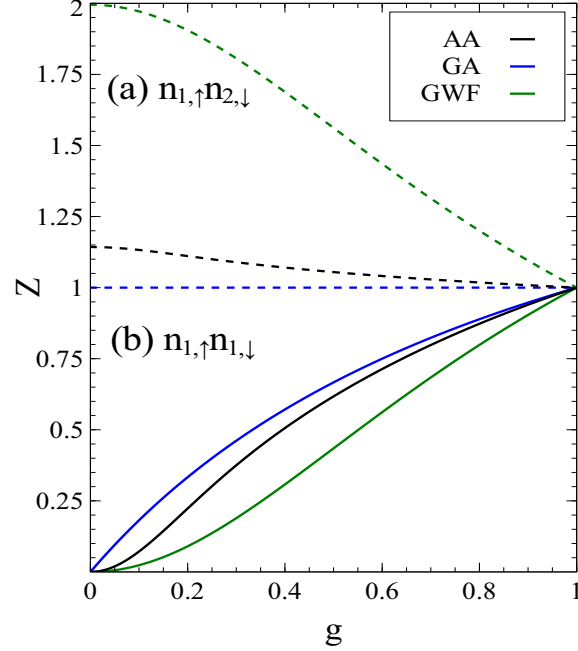
FIG. 2: Shown in the figure are the renormalization factors Z, defined in Eq. 14, as a function of g for the onsite and nearest neighbor density-density interactions evaluated with GWF, AA and GA. The calculation is on a circular chain of 8 Hydrogen atoms described with a STO-3G basis set and separated with an atomic distance of $2.5\mathring{A}$. The noninteracting wavefunction is provided by a restricted Hartree Fock calculation. The optimal Gutzwiller parameter has a value of $g \simeq 0.02$. Energy from AA is about $0.02\,Har$ lower than GWF at $r = 2.5\mathring{A}$ as seen from Fig 1. This energy difference is mainly due to the inaccurate $\langle n_{1,\uparrow} n_{2,\downarrow} \rangle$ evaluated with AA, the current approximation.

onsite double occupancy as well as insufficient enhancement of off-site density-density correlations in the new approximation, as manifested in Fig 2 for an inter-atomic separation of $2.5\mathring{A}$ on the 8 Hydrogen chain system. Consider that the optimal Gutzwiller parameter g is quite small, AA thus gives roughly consistent onsite energy as GWF, but underestimates the inter-site Coulomb repulsion. Consequently, AA gives an underestimated bounding energy. On the other hand, GA deviates more seriously from the benchmark energies. It also gives rise to an unphysical kink close to the bond breaking regime hinting a plausible meta-stable state for the Hydrogen rings.

As the system size increases, renormalization factors, $Z_{\hat{O}}$, in the new approximation gradually approach the GA results, as illustrated in Fig. 3 and Fig. 4 on four typical operators in the half-filled case. Results deviating from half-filling are not shown as both methods are quite close to GWF results(e.g. see Ref [42] for the nonhalf-filling behavior of GA). From Fig. 3 we immediately see noticeable finite size effects for AA. It converges quickly as system size increases. At a system
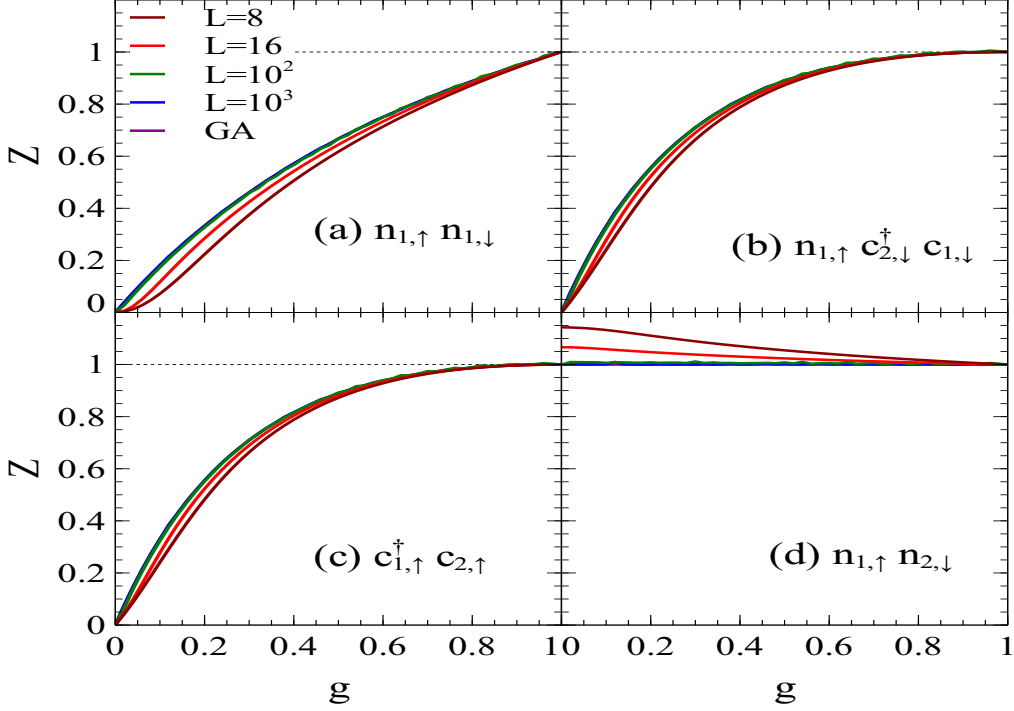
FIG. 3: Renormalization factor Z of four typical operators, $n_{1,\uparrow}n_{1,\downarrow}$, $n_{1,\uparrow}c^\dagger_{2,\downarrow}c_{1,\downarrow}$, $c^\dagger_{1,\uparrow}c_{2,\uparrow}$ and $n_{1,\sigma}n_{2,\sigma'}$ are evaluated with AA for different linear H chain sizes. GA results are also provided for comparison. All graphs share the same legend given in graph (a). Note, no Hamiltonian is needed to scan g dependence of AA and GA, only the Gutzwiller wavefunction is relevant.

size of $N = 10^2$, both AA and GA are nearly indistinguishable from each other. To provide a more quantitative measure on the asymptotic behavior of the new approximation, we present the system size dependence of the relative differences of renormalization factors between AA and GA in Fig. 4. The manifest linear dependence between these two quantities provides solid support on the conclusion that the new approximation asymptotically approaches GA in its infinite lattice limit. This might also suggest that the single band Gutzwiller approximation is quite an unexpectedly stable limit for the Gutzwiller wavefunction formalism with a site-wise local Gutzwiller projector.

AN ATTEMPT TO APPLY THE NEW APPROXIMATION TO MULTIBAND SYSTEMS

Thus far, we have applied the new approximation to single band systems and compared it against GA and other benchmark calculations. Encouraged by the limited success, we would like to generalize it to multiband systems which are physically more interesting and relevant to real problems. The simplest way to do so is to rewrite the Gutzwiller projector in terms of multi-orbital
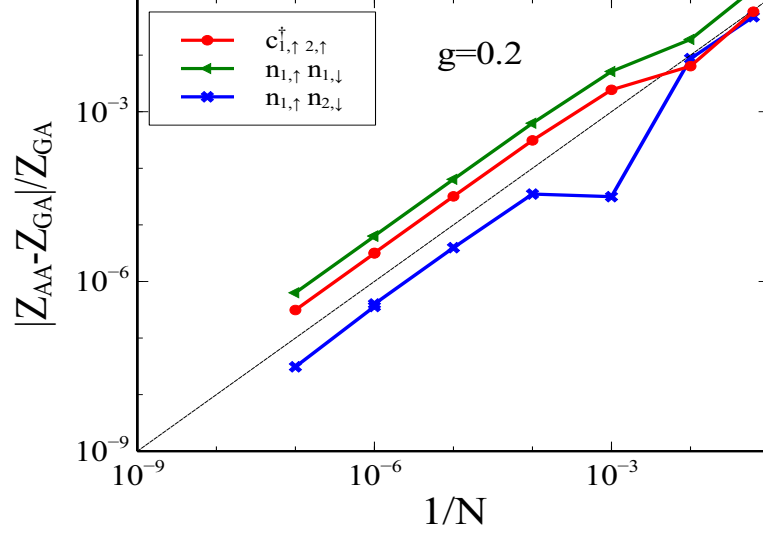
FIG. 4: System size dependence of the relative renormalization factor differences between the new approximation and GA is shown for 3 different operators discussed in Fig. 3. The calculation is carried out with $g = 0.2$. $Z_{n_{1,\uparrow}c^{\dagger}_{1,\downarrow}c_{2,\downarrow}} = Z_{c^{\dagger}_{1,\uparrow}c_{2,\uparrow}}$ for both approximations and is thus suppressed from the plot. Note the log scale is used on both axes.

particle number operators[33], and to follow Eq. 18 to Eq. 23 to impose overall physical constraints. Unfortunately, this is not successful due to the very fact that the proposed approximation itself does not automatically preserve the correct total number of electrons in the system except for the single band case. Indeed, there is no guarantee that relevant physical constraints can be automatically preserved with such a simple and arbitrary wavefunction coefficient averaging scheme. It is a bit of luck that it does work out for single band systems.

However, this is not the excuse for us to give up the whole idea as it does provide us with a candidate recipe to define the renormalization factor for any operator. If we could look into the issue of failure a bit closer and at the same time fill us with confidence that the idea of averaging does work well on single band systems, we might come up with the following interpretation on the validity of the scheme: the averaging procedure does capture part of the correlation physics by giving reduced hopping or onsite double occupancy, etc., but it might describe correlation physics only qualitatively right in generic cases. One way to improve its quality requires more degrees of freedom to enable adjusting the way how the averaging can be carried out. This is actually quite similar to the logic Bunemann et. al. followed in deriving the multiband Gutzwiller approximation[33][32]. There, a set of orbital-wise fugacity parameters are introduced into the Gutzwiller trial wavefunction as adjustable parameters to ensure that the orbital occupations take the values evaluated with

the noninteracting wavefunction[33], or fulfill some constraints of similar nature in case the local Gutzwiller projector is more general than involving only particle operators[36][38]. Their purpose to choose these constraints is for mathematical convenience to enable a rigorous Feynmann diagrammatic expansion to the total energy expectation value evaluated with the Gutzwiller trial wavefunction. Similarly, we will introduce the fugacity parameters into the trial wavefunction and impose other more reasonable physical constraints than creating a connection with information from the noninteracting wavefunction. We will use these fugacity parameters to calculate the renormalization factors defined through averaging. From a statistical point of view, these fugacity parameters define a weight on each wavefunction coefficient. Thus the renormalization factors are defined not as a simple averaging, but as a weighted averaging instead. With a careful choice of weights, a concept readily borrowed from the statistics theory, more reasonable renormalization factors are anticipated to make this new approximation useful. Specifically, we introduce into the trial wavefunction of Eq. 2 a local weighting operator

$$\hat{W}_i = \prod_s (\eta_s)^{\sum_\sigma \hat{n}_{is,\sigma}} \tag{31}$$

where i is site index, s denotes each local correlated orbital and η_s is the weight of that state. Now the Gutzwiller trial wavefunction to be used in defining the renormalization factors is

$$\left|\tilde{\Psi}\right\rangle_G = \left(\prod_i \hat{W}_i \hat{P}_{i,G}\right) |\Psi_0\rangle \tag{32}$$

Note, this wavefunction is not supposed to be used to evaluate physical expectation values of any operator. They are evaluated with Eq. 2 directly, and are thus unique given a Gutzwiller parameter g. Of course, the choice of the weighting operator, $\hat{W}_i$, is very flexible, depending on what kind of physical properties one decides to preserve. In case of a Hydrogen dimer described by a large basis set of 6-311G and to be studied here, we use Eq. 31 to ensure correct total number of electrons in the dimer system. The correlated orbital is chosen to be $1s$ while other orbitals are treated uncorrelated. With these enhanced definitions, the whole averaging formalism can be written out and η_s can be determined analytically as a function of g. Details of the calculation are given in Appendix B. Rather unexpectedly, such a slightly modified averaging scheme gives an exact description on the multiband Hydrogen dimer system within a Gutzwiller trial wavefunction. Besides the nice agreement, one might appreciate a piece of physics embedded in the current approximation. We found that the fugacity parameter has two solutions. One solution gives orbital occupations identical to those of the underlying noninteracting wavefunction, a scenario adopted in the multiband GA formalism, and the other gives the correct orbital occupation in

the Gutzwiller trial wavefunction. The coincidence of the solutions to the assumption underlying the multiband GA formalism might hint that the weighted version of the new approximation is intimately related with the multiband GA formalism that Bunemann et. al derived, just like AA to GA in the single band case. This fact might also speak loud that the new approximation introduced in this work could perform better than GA in capturing the strong correlation physics inherent in the Gutzwiller wavefunction formalism.

CONCLUSION

In sum, we have introduced a very simple yet effective approximation towards the Gutzwiller wavefunction formalism. The simplicity is easily seen by noting that the renormalization factor of an operator is obtained through a direct average over nonvanishing coefficients of a Gutzwiller wavefunction constructed with a density-density type Gutzwiller projector. The effectiveness is supported by the agreement between the new approximation and GA, a well-studied approximation towards the Gutzwiller wavefunction formalism, in the infinite lattice limit of single band systems. Thus the current approximation provides a new perspective towards GA and its underlying assumptions. The proof showing their mutual relationship in Appendix A clearly reveals the grand canonical ensemble nature of GA, which readily prevents it from being applied to finite systems. This very nature of GA also shows itself in the counting argument interpretation of GA[32] but is not clear from the diagram based formulation[33]. In this work, several numerical instances are provided to compare the performance of the new approximation and GA, and to show the asymptotic agreement between these two schemes.

Although it shows an improved performance than GA in single band systems, the naive averaging, however, does not capture the double occupancy and other subtle physical quantities well enough. It also leads to the failure to reach the correct local orbital occupation in a Hydrogen dimer described with a large basis set, the simplest multiband system to be studied. All these can be improved however, at least partially, by introducing more fugacity parameters and physical constraints into the scheme. For example, a simple modification of introducing a weighting operator like Eq. 31 nicely fixes the multi-orbital Hydrogen dimer problem. Actually, such a simple problem indicates a possible close relation between the weighted averaging scheme and the multiband GA formalism. With a more carefully designed weighting factor and more physical constraints included in the formalism, the performance of the current approximation could be systematically improved to reach that of the Gutzwiller wavefunction formalism, with which the current approximation is

built to match.

This new idea is applied most conveniently towards Gutzwiller trial wavefunction with its Gutzwiller projector commuting with number operators. Such a density-density type Gutzwiller projector (d-GPJ) might sound inferior than a more generic Gutzwiller projector (g-GPJ) defined with a general interaction operator to describe local correlation effects[36]. This might cast doubt on the ultimate usefulness of the current scheme to describe real systems. We have to admit that more variational degrees of freedom introduced in g-GPJ are able to give ground state energy. But we also believe that d-GPJ might perform with close quality as g-GPJ. It might also have an additional advantage of being more convenient in its implementation in practice. To fill us with some confidence first, we might find a successful application of d-GPJ in capturing the correct physics. Actually, such a Gutzwiller trial wavefunction has been applied to the Hubbard model and leads to the well-known t-J Hamiltonian in the strong correlation limit[29]. The t-J Hamiltonian, as well as the Heisenberg Hamiltonian as its special case, is the main working horse towards strong correlation limit and spin dynamics. Second, we believe that g-GPJ can only be treated in the multiband GA level in practice. Thus a much fairer comparison would be between the multiband GA and the current scheme, to compare which one is able to perform better. Then, the answer is clear that no party wins over the other for sure. One motivation to develop the formalism involving g-GPJ might be because GA from d-GPJ has to assume $\langle \hat{n}_{(\alpha,\sigma)} \rangle = n^0_{(\alpha,\sigma)}$, a serious limitation in applying GA to multiband systems[33]. While GA from g-GPJ has this constraint removed, it still adopts constraints of similar nature in its derivation in order to carry out a Feynmann diagrammatic expansion[36]. Thus, how far multiband GA can be free from this constraint is unknown to us. In the new approximation introduced in this work, however, local orbital renormalization inherently exists which necessarily gives a local orbital occupation differing from its noninteracting value. One might systematically introduce more and more physical constraints into the formalism to help guide its outcome to be more and more physical. One more advantage of the current scheme against GA is that it is readily applied to Gutzwiller projectors with nonlocal density-density correlations. This fact is quite easy to notice although its practical implementation might be an issue.

APPENDIX A: FROM THE NEW APPROXIMATION TO THE GUTZWILLER APPROXIMATION ON A SINGLE BAND SYSTEM

To establish a connection between these two approximations, let's consider the renormalization coefficient for $\hat{O}(i, j) = c_{i,\sigma}^\dagger c_{j,\sigma}$ with site indices $i \neq j$. Here σ denotes a spin-orbital composite state. Following the spirit leading to Eq. 18 to Eq. 20, one can write down $\mathcal{F}$ and $\mathcal{B}$ for $c_{i,\sigma}^\dagger c_{j,\sigma}$ as

$$\mathcal{F} = \frac{\sum_{\{\boldsymbol{\Gamma}\}} \sum_{\{\boldsymbol{\Gamma}'\}} \left(\prod_{k \neq i,j} g_{\Gamma_k}^2 \right) \left(g_{\Gamma_i} g_{\Gamma_j} g_{\Gamma_i'} g_{\Gamma_j'} \right) \mathcal{S}\left(\hat{O}; \boldsymbol{\Gamma}, \boldsymbol{\Gamma}' \right) \delta_{\boldsymbol{\Gamma}}^{(3)} \delta_{\boldsymbol{\Gamma}'}^{(3)}}{\sum_{\{\boldsymbol{\Gamma}\}} \sum_{\{\boldsymbol{\Gamma}'\}} \mathcal{S}\left(\hat{O}; \boldsymbol{\Gamma}, \boldsymbol{\Gamma}' \right) \delta_{\boldsymbol{\Gamma}}^{(3)} \delta_{\boldsymbol{\Gamma}'}^{(3)}} \tag{33}$$

$$\mathcal{B} = \frac{\sum_{\{\boldsymbol{\Gamma}\}} \left(\prod_k g_{\Gamma_k}^2 \right) \delta_{\boldsymbol{\Gamma}}^{(3)}}{\sum_{\{\boldsymbol{\Gamma}\}} \delta_{\boldsymbol{\Gamma}}^{(3)}} \tag{34}$$

Here $\{\boldsymbol{\Gamma}\}$ and $\{\boldsymbol{\Gamma}'\}$ denote a complete set of occupation configuration on the whole lattice, g_{Γ_i} denotes the g factor at site i with local Fock state Γ_i. In order to reach GA, let's first relax the δ constraints, and note the following relations hold

$$\sum_{\{\boldsymbol{\Gamma}\}} \left(\prod_k g_{\Gamma_k}^2 \right) = \left(\sum_\Gamma g_\Gamma^2 \right)^N \tag{35}$$

$$\sum_{\{\boldsymbol{\Gamma}\}} 1 = \mathcal{D}^N \tag{36}$$

with N total number of lattice sites and $\mathcal{D}$ dimensionality of the local Fock space. We thus have

$$\eta = \frac{\mathcal{F}}{\mathcal{B}} = \frac{4}{\left(\sum_\Gamma g_\Gamma^2 \right)^2} \left(\sum_{\Gamma_i, \Gamma_j, \Gamma_i', \Gamma_j'} g_{\Gamma_i} g_{\Gamma_j} g_{\Gamma_i'} g_{\Gamma_j'} \left| \langle \Gamma_i | c_{i,\sigma}^\dagger | \Gamma_i' \rangle \right| \left| \langle \Gamma_j | c_{j,\sigma} | \Gamma_j' \rangle \right| \right) \tag{37}$$

This expression for η is only intermediate and not fully consistent, as one can see that η fails to vanish as variational parameters g in the standard Gutzwiller projector approaches 0.

Now let us introduce another approximation that the probability of finding a given Fock space occupation is a product of probabilities of finding the specific occupation configuration on each lattice site, or namely,

$$|\langle \{\Gamma\} | \Psi \rangle|^2 = \prod_k p_{\Gamma_k} \tag{38}$$

$$|\langle \{\Gamma_k\} | \Psi_0 \rangle|^2 = \prod_k p_{\Gamma_k}^0 \tag{39}$$

where $p_{\Gamma_k}, p_{\Gamma_k}^0$ denote the probability to find a Fock state Γ_k on site k for the trial and noninteracting wavefunction respectively. This is too big a step forward as $|\langle \{\Gamma\} | \Psi \rangle|^2$ is normally not decomposeable on lattice sites. This approximation can be validated only in the infinite spatial

dimension limit[43], which also underlies the Gutzwiller approximation. With this approximation, it is reasonable to assume

$$g_{\Gamma_k} = \sqrt{\frac{p_{\Gamma_k}}{p_{\Gamma_k}^0}} \tag{40}$$

for the Gutzwiller trial wavefunction defined in Eq. 2 by noting that the local projection operator is site-wise. Feed Eq. 40 into Eq. 37 and we get

$$\eta = \bar{z}_i \bar{z}_j \tag{41}$$

with

$$\bar{z}_i = \frac{2}{\left(\sum_\Gamma p_\Gamma / p_\Gamma^0\right)} \sum_{\Gamma_i, \Gamma_i'} \sqrt{\frac{p_{\Gamma_i} p_{\Gamma_i'}}{p_{\Gamma_i}^0 p_{\Gamma_i'}^0}} \left| \langle \Gamma_i | c_{i,\sigma} | \Gamma_i' \rangle \right| \tag{42}$$

A mean-field type approximation is now introduced for terms involving p_Γ^0 by replacing those terms with their averages. Again, in the infinite spatial dimension limit and under some convenience conditions, p_Γ^0 can be expressed in terms of local occupancy, n_σ^0 on each local state σ, as

$$p_\Gamma^0 = \prod_{\sigma \in \Gamma} n_\sigma^0 \prod_{\sigma \notin \Gamma} \left(1 - n_\sigma^0\right) \tag{43}$$

Then there are two averages needed to be calculated,

$$\overline{\sqrt{p^0 p^0}} = \frac{1}{\left(\frac{\mathcal{D}}{2}\right)} \sum_{\Gamma_i, \Gamma_i'} \sqrt{p_{\Gamma_i}^0 p_{\Gamma_i'}^0} \left| \langle \Gamma_i | c_{i,\sigma} | \Gamma_i' \rangle \right| \tag{44}$$

$$= \frac{1}{\left(\frac{\mathcal{D}}{2}\right)} \sum_{\sigma \notin \Gamma_i'} \sqrt{p_{\Gamma_i}^0 p_{\Gamma_i \cup \sigma}^0} \tag{45}$$

$$= \frac{2}{\mathcal{D}} \sqrt{n_\sigma^0 \left(1 - n_\sigma^0\right)} \tag{46}$$

and

$$\overline{p^0} = \frac{1}{\mathcal{D}} \sum_\Gamma p_\Gamma^0 = \frac{1}{\mathcal{D}} \tag{47}$$

Here $\mathcal{D}$ is the dimensionality of the local Fock space. The prefactor 2 in Eq. 46 accounts for the fact that Γ_i must not contain state σ in it. Plug Eq. 46 and Eq. 47 back to Eq. 42 and one recovers the standard GA definition of z_i given in Eq. 5.

APPENDIX B: THE NEW APPROXIMATION WITH WEIGHTED AVERAGE APPLIED ON THE MULTIBAND H DIMER SYSTEM

To make things simple, let's consider a local basis set composed of one correlated orbital, denoted as s, and $\mathfrak{N}$ uncorrelated orbitals to describe the H-dimer, all defined as Wannier functions such

that they are orthogonal to each other within and between sites. A molecular orbital is thus generically created via operator

$$a_\sigma^\dagger = \sum_i h_{is} c_{is,\sigma}^\dagger + \sum_{i,\alpha \neq s} h_{i\alpha} c_{i\alpha,\sigma}^\dagger \tag{48}$$

where $c_{i\alpha,\sigma}^\dagger$ creates an electron at site i and orbital α with spin σ, or, $c_{i\alpha,\sigma}^\dagger |0\rangle = |i\alpha, \sigma\rangle$. The coefficients satisfy the normalization condition

$$|h_{1s}|^2 + \sum_{\alpha \neq s} |h_{1\alpha}|^2 = \frac{1}{2} \tag{49}$$

with translational invariance implicitly assumed. That the cross terms contributing to the normalization vanish comes from the fact that each atomic orbital is a Wannier function, as mentioned at the beginning of this appendix. Eq. 49 also expresses the electron conservation condition, ensuring each site has half an electron with a specific spin. The noninteracting Hartree-Fock wavefunction can be expressed as

$$|\Psi_0\rangle = a_\uparrow^\dagger a_\downarrow^\dagger |0\rangle \tag{50}$$

for the H dimer, and the Gutzwiller wavefunction is defined as

$$\begin{aligned}
|\Psi\rangle &= g^{\sum_i \hat{n}_{is\uparrow} \hat{n}_{is\downarrow}} |\Psi_0\rangle \\
&= (g-1) |h_{1s}|^2 |1s \uparrow, 1s \downarrow\rangle + (g-1) |h_{2s}|^2 |2s \uparrow, 2s \downarrow\rangle + |\Psi_0\rangle
\end{aligned} \tag{51}$$

with g the Gutzwiller parameter. Expectation value w.r.t the Gutzwiller wavefunction for any operator can be straightforwardly evaluated. Specifically, for particle occupations, there are

$$\langle n_{1s,\sigma}\rangle = \frac{(g^2-1)|h_{1s}|^2 + 1}{2(g^2-1)|h_{1s}|^4 + 1} |h_{1s}|^2 \tag{52}$$

$$\langle n_{1\alpha,\sigma}\rangle = \frac{1}{2(g^2-1)|h_{1s}|^4 + 1} |h_{1\alpha}|^2 \quad \text{for } \alpha \neq s \tag{53}$$

Obviously, it is easy to test that they satisfy the electron conservation condition, Eq. 49, with help of that constraint.

For the current approximation with weighted average enhancement, the local weighting operator is chosen as

$$\hat{W}_i = \eta^{\sum_\sigma \hat{n}_{is,\sigma}}$$

The expectation value for an operator is considered from its denominator, wavefunction normalization, and its numerator respectively. For the normalization factor $\langle \Psi | \Psi \rangle$, there is

$$\langle \Psi | \Psi \rangle = \langle \Psi_0 | \eta^{2\sum_{\sigma=\uparrow,\downarrow} \hat{n}_{is,\sigma}} g^{2\sum_i \hat{n}_{is\uparrow}\hat{n}_{is\downarrow}} | \Psi_0 \rangle$$

$$\simeq \frac{2\eta^4 g^2 + 2\eta^4 + 4C_{2\mathfrak{N}}^1 \eta^2 + \left(2C_{2\mathfrak{N}}^2 + 2\mathfrak{N}\right)}{2\eta^4 + 2\eta^4 + 4C_{2\mathfrak{N}}^1 \eta^2 + \left(2C_{2\mathfrak{N}}^2 + 2\mathfrak{N}\right)} \langle \Psi_0 | \Psi_0 \rangle \tag{54}$$

with C_n^m the usual combinatorial number choosing m elements out of n elements. In the numerator of the renormalization factor in Eq. 54, each term has clear physical interpretation. $\eta^4 g^2$ corresponds to two electrons occupying s orbitals on the same site, η^4 has the two electrons take s orbitals on different sites, η^2 is related to Fock states with only one electron in s orbitals, while the last term corresponds to Fock states with no s orbital. The terms in the denominator of the above prefactor are obtained by ignoring the variational parameter g. What is left acts as the weight to each term in the numerator, a necessary step to normalize a weighted average. Similarly, one can write down the numerator of the expectation value of an operator. For electron occupations, they are

$$\langle \Psi | \hat{n}_{1s,\uparrow} | \Psi \rangle \simeq \frac{\eta^4 g^2 + \eta^4 + C_{2\mathfrak{N}}^1 \eta^2}{2\eta^4 + C_{2\mathfrak{N}}^1 \eta^2} |h_{1s}|^2 \tag{55}$$

$$\langle \Psi | \hat{n}_{1\alpha,\uparrow} | \Psi \rangle \simeq |h_{1\alpha}|^2 \text{ for } \alpha \neq s \tag{56}$$

The constraint of conserved electron occupation on a H dimer requires

$$\langle \Psi | \hat{n}_{1s,\uparrow} | \Psi \rangle + \sum_{\alpha \neq s} \langle \Psi | \hat{n}_{1\alpha,\uparrow} | \Psi \rangle = \frac{1}{2} \langle \Psi | \Psi \rangle \tag{57}$$

Feed Eq. 54, Eq. 55 and Eq. 56 into Eq. 57 and solve for η, and one ends up with two solutions

$$\eta^2 = 0 \tag{58}$$

$$\eta^2 = \frac{\mathfrak{N} |h_{1s}|^2}{\sum_{\alpha \neq s} |h_{1\alpha}|^2} \tag{59}$$

Interestingly, both these two solutions have clear physical interpretations. $\eta = 0$ corresponds to the case where the correlated system has the same local orbital occupation as the underlying non-interacting wavefunction, which is what Bunemann's multiband Gutzwiller approximation starts with. The nontrivial solution of Eq. 59 to η gives the correct charge occupation as the rigorous Gutzwiller wavefunction. One can readily verify this fact by inserting the solution Eq. 59 back to the expressions for $\langle n_{1s,\uparrow} \rangle$ and $\langle n_{1\alpha,\uparrow} \rangle$ and take the constraint of Eq. 49. Actually, one can further verify that this nontrivial solution renders correct expressions for any one and two body operators of the H dimer system.

[1] P. Fazekas, *Lecture Notes on Electron Corrrelation and Magnetism*, 2nd ed. (World Scientific, Singapore, 2003)

[2] P. W. Anderson, *The Theory of Superconductivity in the High-Tc Cuprate Superconductors*, 1st ed. (Princeton University Press, New Jersey, USA, 1997)

[3] P. Coleman, *Heavy Fermions: electrons at the edge of magnetism, in Handbook of Magnetism and Advanced Magnetic Materials, Vol 1: Fundamentals and Theory, ed. by H. Kronmuller and S. Parkin*, 1st ed. (J. Wiley and Sons, New Jersey, USA, 2007) pp. 95–148

[4] D. N. Basov, R. D. Averitt, D. van der Marel, M. Dressel, and K. Haule, Rev. Mod. Phys. **83**, 471 (2011)

[5] O. Parcollet, G. Biroli, and G. Kotliar, Phys. Rev. Lett. **92**, 226402 (2004)

[6] J. Liu, J. Schmalian, and N. Trivedi, Phys. Rev. Lett. **94**, 127003 (2005)

[7] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many Particle Systems*, 1st ed. (Dover Publications, New York, US, 2003)

[8] D. P. A. Chubukov and J. Schmalian, *A Spin Fluctuation Model for d-wave Superconductors, in The Physics of Superconductors, ed. K.H. Benneman and J. B. Ketterson*, 2nd ed. (Springer, Berlin, Germany, 2008) pp. 1349–1407

[9] N. E. Bickers and D. J. Scalapino, Ann. Phys. **193**, 206 (1989)

[10] C. X. Chen and N. E. Bickers, Solid State Commun. **82**, 311 (1992)

[11] N. E. Bickers, *in Theoretical Methods for Strongly Correlated Electrons, edited by D. Senechal, A. Tremblay, and C. Bour-bonnais* (Springer-Verlag, New York, USA, 2004) p. 237

[12] J. C. Bonner and M. E. Fisher, Phys. Rev. **135**, A640 (1964)

[13] J. E. Hirsch, R. L. Sugar, D. J. Scalapino, and R. Blankenbecler, Phys. Rev. B **26**, 5033 (1982)

[14] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)

[15] G. Kotliar, S. Y. Savrasov, G. Pálsson, and G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)

[16] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)

[17] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975)

[18] R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008)

[19] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992)

[20] J. Zeng, *Quantum Mechanics, Volume I*, 3rd ed. (Science Press, Beijing, China, 2000)

[21] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (Nov 1964), http://link.aps.org/doi/10.1103/PhysRev.136.B864

[22] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (Nov 1965), http://link.aps.org/doi/10.1103/PhysRev.140.A1133

[23] P. W. Anderson, G. Baskaran, Z. Zou, and T. Hsu, Phys. Rev. Lett. **58**, 2790 (1987)

[24] P. W. Anderson, Science **235**, 1196 (1987)

[25] M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)

[26] J. Hubbard, Proc. Roy. Soc. A **276**, 238 (1963)

[27] M. C. Gutzwiller, Phys. Rev. **134**, A923 (1964)

[28] J. Spałek, Phys. Rev. B **37**, 533 (1988)

[29] e. a. A. P. Balachandran, *Hubbard model and anyon superconductivity* (World Scientific Publishing Co., Singapore, 1990)

[30] D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984)

[31] K. K. Tohru Ogawa and T. Matsubara, Prog. Theor. Phys. **53**, 614 (1975)

[32] J. Bumenann, Eur. Phys. J. B **4**, 29 (1998)

[33] J. Bunemann, F. Gebhard, and W. Weber, Journal of Physics: Condensed Matter **9**, 7343 (1997)

[34] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)

[35] W. Metzner and D. Vollhardt, Phys. Rev. B **37**, 7382 (1988)

[36] J. Bünemann, W. Weber, and F. Gebhard, Phys. Rev. B **57**, 6896 (1998)

[37] T. Okabe, J. Phys. Soc. Japan **65**, 1056 (1996)

[38] M. Fabrizio, Phys. Rev. B **76**, 165110 (2007)

[39] W. F. Brinkman and T. M. Rice, Phys. Rev. B **2**, 4302 (1970)

[40] V. Anisimov, I. Nekrasov, D. Kondakov, T. Rice, and M. Sigrist, The European Physical Journal B - Condensed Matter and Complex Systems **25**, 191 (2002), ISSN 1434-6028, http://dx.doi.org/10.1140/epjb/e20020021

[41] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*, 1st ed. (Cambridge University Press, Cambridge, UK, 2009)

[42] W. Metzner and D. Vollhardt, Phys. Rev. B **37**, 7382 (1988)

[43] This statement can be formally argued in the following way. Introduce a bosonic operator $\hat{B}_\Gamma^\dagger = \prod_{\sigma \in \Gamma} \hat{n}_\sigma \prod_{\sigma \notin \Gamma} (1 - \hat{n}_\sigma)$ for projecting out a Fock state Γ on each lattice site and the occupation probability for a given occupation configuration $\{\Gamma_i\}$ can be expressed as $P = \langle \Psi | \prod_i \hat{B}_{i,\Gamma_i}^\dagger | \Psi \rangle$. This expression is then decomposeable in site indices in the infinite dimension limit.